Revisit Behavior in Social Media:
The Phoenix-R Model and Discoveries

Flavio Figueiredo[1], Jussara M. Almeida[1],
Yasuko Matsubara[2,3], Bruno Ribeiro[3], and Christos Faloutsos[3]

[1] Department of Computer Science, Universidade Federal de Minas Gerais
[2] Department of Computer Science, Kumamoto University
[3] Department of Computer Science, Carnegie Mellon University

Abstract. How many listens will an artist receive on a online radio? How about plays on a YouTube video? How many of these visits are new or returning users? Modeling and mining popularity dynamics of social activity has important implications for researchers, content creators and providers. We here investigate the effect of revisits (successive visits from a single user) on content popularity. Using four datasets of social activity, with up to tens of millions media objects (e.g., YouTube videos, Twitter hashtags or LastFM artists), we show the effect of revisits in the popularity evolution of such objects. Secondly, we propose the PHOENIX-R model which captures the popularity dynamics of individual objects. PHOENIX-R has the desired properties of being: (1) parsimonious, being based on the minimum description length principle, and achieving lower root mean squared error than state-of-the-art baselines; (2) applicable, the model is effective for predicting future popularity values of objects.

1 Introduction

How do we quantify the popularity of a piece of content in social media applications? Should we consider only the audience (unique visitors) or include revisits as well? Can the revisit activity be explored to create more realistic popularity evolution models? These are important questions in the study of social media popularity. In this paper, we take the first step towards answering them based on four large traces of user activity collected from different social media applications: Twitter, LastFM, and YouTube[4].

Understanding the popularity dynamics of online content is both a challenging task, due to the vast amount and variability of content available, as it can also provide invaluable insights into the behaviors of human consumption [6] and into more effective engineering strategies for online services. A large body of previous work has investigated the popularity dynamics of social media content, focusing mostly on modeling and predicting the *total number of accesses* a piece of content receives [5, 6, 9, 17, 21].

However, a key aspect that has not been explored by most previous work is the effect of revisits on content. The distinction between audience (unique users), revisits (returning users), and popularity (the sum of the previous two) can have large implications for different stakeholders of these applications - from content providers to content

---

[4] http://twitter.com http://lastfm.com http://youtube.com